Articles

Qualitative Methods: What Are They and Why Use Them?

Shoshanna Sofaer

Objective. To provide an overview of reasons why qualitative methods have been used and can be used in health services and health policy research, to describe a range of specific methods, and to give examples of their application.

Data Sources. Classic and contemporary descriptions of the underpinnings and applications of qualitative research methods and studies that have used such methods to examine important health services and health policy issues.

Principal Findings. Qualitative research methods are valuable in providing rich descriptions of complex phenomena; tracking unique or unexpected events; illuminating the experience and interpretation of events by actors with widely differing stakes and roles; giving voice to those whose views are rarely heard; conducting initial explorations to develop theories and to generate and even test hypotheses; and moving toward explanations. Qualitative and quantitative methods can be complementary, used in sequence or in tandem. The best qualitative research is systematic and rigorous, and it seeks to reduce bias and error and to identify evidence that disconfirms initial or emergent hypotheses.

Conclusions. Qualitative methods have much to contribute to health services and health policy research, especially as such research deals with rapid change and develops a more fully integrated theory base and research agenda. However, the field must build on the best traditions and techniques of qualitative methods and must recognize that special training and experience are essential to the application of these methods.

Key Words. Qualitative research methods, case study research, key informant interviews, focus group research, participant observation

Behind every quantity there must lie a quality.

Gertrude Jaeger Selznick, Ph.D.

The field of health services research is young. Its multi-disciplinary practitioners have borrowed conceptual frameworks and data collection and analytic methods from a wide range of social and behavioral sciences, as well as

from public health and medicine. Many health services and health policy researchers have used qualitative methods, either alone or in combination with quantitative approaches. As the epigram (from a lecture by a revered epistemologist, sociologist, and survey researcher) indicates, quantification, which facilitates access to powerful statistical tools, must rest on prior work on conceptualization and operationalization, and on valid and reliable measurement (e.g., see Babbie 1998). Qualitative methods have the potential to contribute significantly to the development of meaningful "quantities"; however, they have inherent as well as instrumental value. Some phenomena, including historical events, are so unusual that by the time one has a way to quantify them, they have either changed or disappeared.[1] Quantification is an excellent way to "slice up" phenomena so that they become manageable and discrete elements of an overall conceptual framework or analytic plan. Everything has the defects of its qualities, however, and the defect of quantification is that it does not always support, as well as qualitative work, the understanding of complex, dynamic, and multi-dimensional "wholes" (Patton 1975).

In addition, the state-of-the-art of quantification is mixed across potentially important independent and dependent variables. If we focus research only on what we already know how to quantify, indeed only on that which can ultimately be reliably quantified, we risk ignoring factors that are more significant in explaining important realities and relationships. Qualitative methods help provide rich descriptions of phenomena. They enhance understanding of the context of events as well as the events themselves. The use of these methods tends to enhance peripheral vision, which is especially important at the early stages of inquiry, as noted further on. In addition, qualitative methods can indeed help to identify patterns and configurations among variables and to make distinctions. Thus, qualitative research not only serves the desire to describe; it also helps move inquiry toward more meaningful explanations.

For all of these reasons, the use of qualitative methods is growing in health services and health policy research (Bowling 1997). Unfortunately, however, we need to appreciate more fully the differences between systematic, rigorous, well-designed qualitative research and well-intentioned but poorly implemented attempts to supplement quantification with more

open-ended interviews. Paradoxically, many decry qualitative methods as inherently biased because these methods depend so much on the perspective and skills of the researcher, while others assume that qualitative research can be carried out by anyone regardless of their training, knowledge, and experience.

This article explores the reasons why and the situations where it is appropriate to use qualitative methods, and it discusses some common uses of these methods in health services and health policy research. It presents a spectrum of qualitative methods that reflect the various ways in which they can be used. It closes with a few examples of qualitative research to illustrate the points.

WHY USE QUALITATIVE METHODS?

It can be useful to conceptualize research as the process of reducing our uncertainty about important phenomena or questions. This implies that the development of knowledge involves the gradual reduction of uncertainty. If we adopt such a developmental approach to inquiry, it becomes evident that, at the outset, there is uncertainty not only about answers, but about what the right questions might be; about how they should be framed to get meaningful answers; and about where and to whom questions should be addressed. As understanding increases, some of the right questions emerge, but uncertainty remains about whether all of the right questions have been identified. Further along, confidence grows that almost all of the important questions have been identified and perhaps framed in more specific terms, but uncertainty still exists about the range of possible answers to those questions. Eventually, a high level of certainty is reached about the range of almost all of the possible answers. In many (but probably not all) cases, therefore, as inquiry proceeds, it can move from being quite unstructured and probably largely qualitative in nature, to being quite structured and probably largely quantitative in nature. We can, as Abraham Kaplan puts it (1964), move from the context of discovery to the context of justification.

One way to make this more concrete is to put it in terms of how to ask questions. In this developmental framework, initial questions are very open-ended; often the researcher just asks what appear to be relevant individuals to describe, in their own way, their experiences and responses concerning a given situation or issue. Over time, questions remain open-ended but become more specific, either in their initial wording or in accompanying "probes."

As we move to more quantitative research, questions become closed-ended, that is, they have specified response options; in addition, the sequence of questions has become predetermined rather than left to the discretion of the researcher. As Patton has noted (1990: ch. 7), there is a discipline, as well as an art, to asking truly open-ended questions: genuine awareness of whether you are asking an open-ended question can be viewed as one of the core disciplines of the qualitative researcher.

Another way to shed light on reasons to use qualitative methods is to address the issue of theory. Health services researchers tend to borrow theories from social science disciplines (e.g., economics, psychology, organizational behavior, political science, sociology, and anthropology). They have done little to develop full-scale, independent theories. It can be argued that the field has reached the point at which it has become most valuable to pay greater attention to the construction of theories that integrate understandings from our multiple disciplinary roots. A shared set of concepts or constructs derived from more integrative and generic theories could help ensure that key questions are addressed and that key variables are measured, and measured consistently, across studies in widely different settings and over time. This would contribute greatly to our ability to build a coherent body of knowledge with perhaps greater staying power.

Qualitative methods are quite useful in constructing or developing theories or conceptual frameworks or, to put it another way, in generating hypotheses. These methods can also be used in refining theories and hypotheses through preliminary testing. There is an important parallel to be drawn here between rigorous qualitative and rigorous quantitative research. In quantitative research, building on epistemological assumptions regarding whether it is ever possible to confirm (rather than disconfirm) a hypothesis, researchers typically try to reject a null hypothesis. An important insight from Patton is that when qualitative researchers begin to see a pattern that might be articulated as a hypothesis, they should then search systematically for evidence that will lead to the rejection of the hypothesis (Patton 1990: ch. 8). When qualitative researchers embrace this level of discipline, they take important and concrete steps to protect against bias, whether intentional or unintentional.

Although qualitative methods are more frequently used in theory development and refinement, some argue that they can also be used in testing theory. No less apparently quantitative a social scientist than Donald Campbell has argued that qualitative methods can be used to collect the data needed for experimental studies (1975). Some qualitative researchers

may argue that this is an inappropriate use of their approach, because it involves a significant shift away from the type of research that involves, as Charles Ragin (1994) has noted, in-depth, multi-dimensional exploration of a limited number of complex wholes. If one views the use of qualitative methods as a continuous rather than a dichotomous variable, however, it is easier to conclude that qualitative methods can inform and be informed by typical practice in quantitative research, and that it can provide ways to make reliable observations that would not otherwise be possible.

USING QUALITATIVE METHODS IN HEALTH SERVICES AND POLICY RESEARCH

In the field of health services research, qualitative methods have been used to describe many kinds of complex settings and complex interactions. These include interactions among patients, families, and clinicians; within, between, and among professional groups and organizations; in communities; and in markets. Classic studies by Anselm Strauss and Barney Glaser, for example, used qualitative methods (what they called "grounded theory" [1967]) to elucidate relationships among professionals and between professionals and patients. Examples of the continuing tradition include Glaser's work (1965) on the issue of whether or not to tell patients they are dying, Strauss's work highlighting the virtually continuous process of negotiation in the specification of roles and relationships in health care settings (1978), and Corbin and Strauss's more recent work on the life experience of family members who care for people with chronic illnesses in home settings (1988).

Qualitative research also plays an important part in clarifying the values, language, and meanings attributed to people who play different roles in organizations and communities. They allow people to speak in their own voice, rather than conforming to categories and terms imposed on them by others. Often, qualitative researchers find that they are giving voice, in particular, to those who are otherwise rarely heard, such as patients or workers far down in the hierarchical chain of command. This latter tradition can be seen as the application of the techniques of anthropology. The research methods of anthropology were designed both to describe and to understand distinctive cultures in far away places. They have proved to be very useful, as well, in understanding groups of people with similarly distinctive cultures who live in our nation, who are often either almost invisible (e.g., patients

in nursing homes or mental hospitals), alienated from mainstream society
(e.g., people with tuberculosis who are resistant to participation in treatment
programs), or simply not familiar to average Americans (e.g., many immi-
grant groups). From Goffman's classic study of mental hospital patients as
"inmates" (1961) to Diamond's more recent study of nursing homes (1992), a
sensitivity to the ways in which realities and roles are socially constructed and
reconstructed has helped illuminate the unspoken assumptions underlying
relationships between "health care providers" and people with relatively little
power in society.

Qualitative methods are also very useful in inquiries into developmental
and historical processes within institutions, communities, and markets. One of
the great advantages of qualitative methods is that they enhance the capacity
not only to describe events but to understand how and why the "same" events
are often interpreted in a different, sometimes even conflicting manner, by
different stakeholders. In the context of policy research in particular, qualita-
tive methods have been used to document the perspectives and interactions
among multiple stakeholders. This is of great value in studies of policymaking,
of policy implementation, and even of policy consequences.

A conundrum is inherent in this extraordinary capacity of qualitative
methods to capture the relative rather than the absolute nature of "truth." Do
we learn more about our informants or do we learn more about the events on
which they are commenting? At least some readers have probably seen the
film *Rashomon,* the directorial debut of the late great Japanese director Akira
Kurosawa. This remarkable work, shot in the 1950s but set centuries ago, tells
and retells the story of a robbery, rape, and murder from the perspective of
multiple participants and observers, including the ghost of the victim. As the
film unfolds, the audience learns a great deal not only about the meaning of
the events to different people, but about the motivations, too, that lead them
to present events as they do. Often, these motivations reflect quite human
desires to maintain what the French call *amour propre,* or what others might
call simply "image." It is not clear that anyone "knows," at the end, what
really happened. Some might argue that any single statement of events would
necessarily be less than complete or accurate. Qualitative researchers often
use the term "triangulation" to describe the process of examining different
perspectives in order to identify at a minimum what all informants *seem* to
agree took place, or what it means (Miles and Huberman 1994; also see Patton
in this issue). This is almost as if we assume that the truth exists only in the
space where multiple Venn diagrams converge. Perhaps it is more honest to

admit that some of the truth may be found in the places in the diagram where the circles do not converge.

Qualitative methods are frequently used in health services and health policy evaluations (Caudle 1994; Scheirer 1994). In some ways, evaluative research is a special kind of social research, in which the dependent variables are by definition normative. In some cases the "desired outcomes" of a program are well known in advance, but in other cases the full range of outcomes (desired and perhaps not so desired) of interventions are not fully known at the outset. This is where the enhanced "peripheral vision" of qualitative methods can be of special use. As an evaluation researcher, however, I have found that the greatest value of qualitative methods in evaluative research has been in providing a far more sophisticated approach to the specification of the "independent variable," that is, the intervention being evaluated. All too often, the independent variable in evaluative research, in health as well as in other fields, has been conceptualized as dichotomous: people either got the "intervention" or they did not. To anyone who has been close to the process of implementing programs, particularly novel programs, this is naïve at best. Over 20 years ago, in compiling a meta-evaluation of substance abuse prevention programs, colleagues examined over 100 studies of different interventions to determine both the rigor of the evaluation methods used and differences in the outcome across categories of intervention (Schaps et al. 1981). In a substantial subset of the studies, not enough data were available on the nature of the intervention to categorize it even at the most rudimentary level. Those studies presented quite a bit of outcome data but almost no process data. In many cases, in sum, it was impossible to know exactly what intervention was being evaluated. In contrast, the "rich description" capacity of qualitative methods can result in a far more complete, and often far more compelling articulation of the intervention, one that can be used both in helping to explain outcomes and in encouraging the adoption of effective practices.

Furthermore, "there is many a slip between cup and lip," and programs are rarely implemented exactly as planned. In multi-site evaluation studies, considerable variation typically occurs in the nature and extent of implementation of the "same" intervention across sites. And the level and nature of implementation can also change over time, either increasing or decreasing—or just swinging back and forth. Qualitative methods have been invaluable in exploring issues such as these: How did those implementing the intervention perceive the intentions of the original planners? Did they have the same agendas or different ones? Did they try to implement the

original plan? If not, why not, and what did they try to implement? What problems did they face in implementation? What new opportunities arose that they wanted to take advantage of? What problems did they solve and how? How did these solutions affect the intervention? Did they try to change the intervention in response to early perceptions of what was and was not working? What changes did they succeed in making? What changes were they unable to make?

Within the field of evaluation, distinctions are often made between formative and summative studies (classic references include Cronbach 1963 and Scriven 1967). Formative studies are designed to provide early feedback to program operators. But process evaluation and formative evaluation often become confused with each other. Although most formative studies pay a lot of attention to issues of process and program implementation, they can, and sometimes do, use both early quantitative data and qualitative methods (such as focus groups and interviews) to explore preliminary outcomes. Summative studies are best conducted when an intervention has stabilized, when its implementation has become as close to complete as it is likely to get, and when the intervention has become consistent over time and place. Qualitative methods can make a contribution not only in identifying the point at which stabilization has occurred (seldom as soon as we think), but also in documenting in detail the "final" intervention and in providing data on both context and outcomes.

Finally, qualitative methods have a clear role in the development of methods and materials. In particular, focus groups and one-to-one cognitive testing have become state-of-the-art tools for the development of reliable and valid survey instruments. Focus groups are used to get an initial sense of the dimensions that are of particular relevance to a topic and set of respondents, so that these dimensions can be included in a survey. Cognitive testing is used to ensure that the items and response options on a survey are interpreted consistently across potential respondents as the survey developers intended (see, for example, the Consumer Assessment of Health Plans Study described by Hurley in this issue). Underlying both of these activities is a fundamental strength of qualitative methods: their ability to explore meanings and, in particular, meanings ascribed to events and circumstances by actors rather than observers.

The same methods are also used to develop the messages and materials used in social marketing campaigns and educational interventions. Focus groups can be of particular assistance in identifying "hot button" concerns of potential audiences and in helping to craft key messages. Cognitive testing

can help determine whether people are in fact "getting the message" from specific materials and products and whether these products help move people to the decisions and actions that the developers have in mind.

A SPECTRUM OF METHODS

A wide spectrum of qualitative research methods can be used, and has been, in health services and health policy research. The "purest" form of qualitative research, with roots in anthropology and sociology, is often known as ethnography or naturalistic inquiry. Naturalistic inquiry most often involves a long-term exposure to a setting and/or a group of people (Lincoln and Guba 1985). The investigator makes extensive use of unstructured observations as well as conversations (they can hardly be called interviews). Usually, detailed field notes are kept of observations and conversations, and in the best research, discipline is used in keeping boundaries set between what is observed and the observer's related responses and interpretations. In some cases, investigators may also gather and analyze the content of a wide range of "documents" (official and unofficial) or other "traces" of phenomena and events. This kind of inquiry is most likely to be used when the level of uncertainty is high (when we are not sure what questions to ask of whom), when there is little or no theory to direct our attention, and when situations are novel or complex.

One subset of this kind of inquiry involves not just observation, but "participant observation," in which the investigator becomes, quite explicitly, a part of the setting or process being examined. For those steeped in the Cartesian duality between observer and observed, this approach is indeed hard to fathom. Yet in many settings, it can be impossible (or as near as makes no difference) to get sufficient exposure to a setting or a group without becoming a participant. For example, many ethnographers assert that it is impossible to understand the life experience of someone in a "total institution," such as a prison or a mental hospital, without not just spending a good deal of time in such an institution but also being subject to the same constraints as those who find themselves forced, by circumstance or the law, to be there. But participant observation does not have to be quite so dramatic. Over 20 years ago, when I was (and looked) a lot younger, I participated in a weekly "support group" for young adults who were at risk for serious drug or alcohol problems. This participant observation was part of an evaluation of an innovative multi-intervention substance abuse prevention program. Because of the ground

rules of the group, a person could not really come to even one meeting without being initiated into membership. The young people were quite comfortable with adults being present, and the reasons for my presence were explained to all of the members. I must admit that I had a good time. More important, I learned far more by attending a few group meetings (about the kinds of young people who participated; about the dynamics of the group; about the style used by the group leaders; and, in fact, about the mismatch between the program designers' expectations and the actuality of the program) than I could have learned by using either more structured qualitative methods, such as interviews, or more closed-ended methods, such as surveys.

A more common set of qualitative methods are those used in case study research (e.g., Yin 1984). Some studies involve a single case, or an event history; others involve multiple cases, chosen either at random or more often purposively to meet the requirements of predetermined typologies. Case study research can also be fairly long-term, but typically it does not involve continuous immersion in a setting or group. Rather, it involves from one to perhaps five short but intensive exposures to a setting or group. In some cases, a single researcher goes on site alone; in other cases a small team of researchers is used. In virtually all instances, case studies use not only purposive sampling of sites, but purposive sampling of informants and experiences.[2] By definition, this means that the investigators have at least some idea of what they are looking for and where they need to go to find it. The important variables have already been identified, and a decision has been made about whether to explore the variable or control for it in the sampling design.

The methods used in case study research are similar to those of naturalistic inquiry, except that data collection is typically more structured. The most common methods used are key informant interviews, structured observations of events and interactions, and the collection and content analysis of relevant documents. Interviews can vary from the unstructured to the highly structured, but the questions remain overwhelmingly open-ended. Typically, a list of questions is specified and there are "probes" associated with certain questions. Just as typically, the field researcher is given considerable latitude about the exact wording of the questions and their sequencing. It is very common for an informant to start answering a question that has not yet been asked; at that point, the experienced researcher exercises judgment about whether to follow the lead of the interviewee or to return to the sequence in the interview protocol. It is also common for case study researchers to use quantitative data. For example, in studying a community, researchers

can and often do take advantage of the available data on the community's demographic characteristics and/or health status. A researcher studying an organization or a market may want to examine data on health care utilization and/or financial information. These data may simply provide background, or they may generate specific questions to ask about the informants' interpretation(s) of certain data or about their perceptions of the impact of the situation as reflected in the data.

Structured observations of meetings and events are probably not used as frequently as are key informant interviews in case study research. Such observations, however, can be invaluable in providing investigators with a direct experience of the ways in which actors interact in a setting. Concerns are often expressed about whether the presence of the researchers influences these interactions so extensively that the interactions are no longer authentic. Many researchers discover that experience teaches greater humility: after a little while, a skilled and disciplined observer does not have that much of an effect on the interactions. Exceptions clearly occur, however, and it can be difficult, especially for the less experienced field researcher, to assess whether he or she has witnessed a naturally occurring event or a well-staged drama.

Content analyses of documents can also vary. In some cases, documents are being examined in order to identify the "facts" of a situation or a series of events. In other cases, again closer to the spirit of naturalistic inquiry, the analysis focuses more on the meanings embedded either in language or, in some cases, in images. This kind of analysis typically involves the identification of assumptions, values, and priorities. It can also help to illuminate differences in the perception of similar events across different actors.

The two last methods have already been mentioned: focus groups (Morgan 1988) and cognitive interviews. These methods are typically used to explore highly specific issues. Nevertheless, they retain the quality that all information-gathering efforts should have: leaving lots of room for investigators to be surprised. The focus group involves bringing together a group of individuals chosen to meet a specific profile of characteristics. Typically, groups are intentionally homogenous along some dimensions and heterogeneous along other dimensions. A structured but still informal setting is used to explore a limited number of "focus questions." Focus groups are best used instead of individual interviews when it is clear that the interactions among group members will be as illuminating as the statements of any individual. As with other qualitative methods, focus groups are often combined with more quantitative approaches such as surveys. Brief surveys are often administered at the end of the group discussion. In addition, recent technological advances

have made it possible for focus group members to respond anonymously to closed-ended questions posed by the group facilitator. In some cases the statistical distribution of responses can be quickly calculated and graphically displayed on a computer screen to become grist for additional discussion. Sometimes responses are tested at multiple points in time to determine whether they are affected by additional information or ideas presented to the group.

Cognitive interviews are a still more specific data collection method. In their typical use in survey development, for example, a small number of one-on-one interviews are conducted with people who would meet the criteria for completing a particular survey. In each interview, the person is asked to complete all or a part of the survey instrument being tested. Sometimes the persons interviewed are asked to "think aloud" as they go along, restating questions (and sometimes response options) in their own words and indicating the cognitive process they use in choosing an answer. At other times the survey is completed first (although the interviewee is sometimes able to ask clarifying questions along the way), and the interviewees are then "de-briefed" to go over their interpretation of the questions and the reasons why they gave the answers they did. This is certainly *not* naturalistic inquiry. But it does reinforce the notion that qualitative methods are excellent at helping us understand how people perceive and interpret language and their own experiences.

EXAMPLES OF THE USE OF QUALITATIVE METHODS

Here are some examples of the use of qualitative methods in my experience. The first is from a study in which my role was relatively minor. The SUPPORT study is a multi-site, and largely quantitative, study of decisions made when a patient in an intensive care unit is, or may be, close to the end of life. In the last phase of the study, nurses were assigned to ICUs to provide fairly unstructured assistance to both family members and health professionals facing difficult decisions about whether and how to intervene. Nurses were asked to select every "nth" patient, whom they "entered" in the study, and to keep a daily journal to record their experiences in working with the patient, the family, and health professionals. The contents of this daily journal were loosely structured into four or five issues. The goal was to collect detailed qualitative data on the interaction of the nurses with others, and on the dynamics among the various "players" in the intense drama of

the end of life, in order to understand better the effects of these dynamics on whether decisions were made and what those were; and by whom, at what point, and with what apparent consequences.

At this stage of inquiry into end-of-life decision making, there was some understanding, but it was far from complete, of who might be involved, in what ways, and when. Little was known about those issues that would be perceived as significant by different actors. It was clear that professional perspectives might well be divergent from those of the patient and the family. Clearly, this study was investigating a highly sensitive situation, and the nurse observer was also an active participant in the drama, reporting not only on what she was (and was not) able to do with respect to engaging people in a dialogue, but also on their reactions to her attempts. These qualitative data are still being fully analyzed, but even preliminary analysis highlighted the particular importance of divergent opinions within families, and between family members and patients, as a barrier to decision making.

A second example is a study under my direction that examined the implementation, at the federal and state levels, of legislation known as the Synar Amendment, after its now sadly deceased sponsor, Congressman Mike Synar of Oklahoma. This amendment required that states pass (if they had not already) a bill prohibiting the sale of tobacco products to those under 18. In addition and more significantly, the amendment required that states document that they were enforcing the legislation. The "stick" in the legislation made it both controversial and interesting: states that did not document meaningful enforcement risked losing, progressively, from 10 to 40 percent of their substance abuse block grant. The goal of the study was to document and perhaps assess the extent of implementation, to identify factors that helped or hindered implementation, and to observe unexpected or at least unintended consequences of the legislation. The study examined implementation both at the federal level and in a purposive sample of states, and it was conducted over several years in order, as originally designed, to see if the "sticks" in the legislation were ever needed and were ever used. The study built on a mix of theories, including interest group theory and interorganizational relations theory, but it sought to refine these theories and extend their application to fairly new ground. A number of potential factors helping or hindering implementation, especially at the state level, could already be identified at least on a preliminary basis. Most of the key stakeholders were already known. Because of this, both states and informants at both the federal and state level could be sampled to address key issues and stakeholders. One particular interest was in how substance abuse professionals and policymakers, who had

not historically interacted with those in the tobacco control world, would deal with the challenge inherent in this new legislation. The methods used included semi-structured key informant interviews, tracking of legislative changes, and content analysis of key documents (including comments made on the Notice of Proposed Rule Making for the amendment). Perhaps the most novel method used was a content analysis of interactions among tobacco control activists who participated in a pre–World Wide Web computerized strategy exchange. Tracking these interactions provided a good deal of information on the activists who were truly engaged in the issue and those who were not. The information proved quite surprising.

The history of this study shows one important value of qualitative methods: their ability to deal with unforeseen events. The implementation of the Synar Amendment was significantly affected by such events, including health care reform efforts (which dominated the focus of many federal health officials); the massive shifts in political power initiated by the election of 1994, at both the federal and state level; and the introduction by the Clinton administration of a far broader and more ambitious youth tobacco initiative spearheaded quite unexpectedly by the Food and Drug Administration. These events created delays in implementation that kept the "outcome" in an ever further retreat, and it required the examination of new issues and new stakeholders. But it also provided an opportunity to generate unexpected insights. For example, a clear difference could be identified between the level of national leadership engagement for the Synar Amendment, a purely congressional initiative, and the FDA Youth Initiative, very much an executive branch initiative closely linked to important broad themes of the Clinton administration. The flexibility of our design permitted us to expand our data collection efforts to new key informants and new documents. Reinterviews of the same informants over time in response to unexpected events also proved illuminating.

The final example also draws from tobacco control research, but it can be seen more broadly as an evaluation of the extent to which multiorganizational coalitions contribute to the passage of desired state policies and the identification of coalition attributes that are most instrumental in their effectiveness. This study, of the Robert Wood Johnson Foundation's SmokeLess States Initiative, was driven by a fairly detailed conceptual framework that drew from multiple theoretical perspectives and the work of several researchers.[3] The conceptual framework articulated a set of intermediate outcome variables for the work of tobacco control coalitions (the passage of key legislation or ordinances, the institutionalization of tobacco control efforts,

and a sustained commitment of participants to the coalition); a set of contextual variables that would mediate coalition effectiveness (e.g., public attitudes toward tobacco control at baseline, community experience in collaborative efforts); and a sequence of independent variables relating to the structure and functioning of the coalitions themselves. After preliminary research, a "theory of action" (Patton 1997) was also developed for the SmokeLess States Initiative that articulated a series of assumptions about the ways in which coalitions would make a difference in tobacco control policy.

The study involved multiple site visits over two to three years to 12 program sites. A multi-disciplinary team of field researchers, proximate to the sites, was used; each site was studied by a single researcher and most researchers were responsible for two sites. This study is an example of a gradual shift from a largely qualitative to a mixed qualitative and quantitative methodology, as learning and understanding grew. It began with, and retained until the end, typical case study research methods described earlier: in-depth interviews with key informants; structured observations of key meetings and other events; and the content analysis of documents.

Even at the outset, given the explicit conceptual frameworks, instruments were fairly structured. Interviews were guided by a modular interview protocol that was actually far too long to be administered at one time to a single informant. Instead, in the course of selecting informants, the field researchers used a reconnaissance technique to identify which persons knew the most about different topics, to avoid wasting time asking questions on which people did not view themselves as knowledgeable. (Used correctly, this approach also helps to ensure that a research team has the right set of informants overall.) Observations were also structured by detailed protocols. Site visit reports were developed and, at the end of the period of observation, integrated case studies were written, using a common format. This format specified clearly a place for the interpretive analysis and comments of the researcher as well as a place for the researcher to assess whether and how she or he might have had an effect on the site over the course of the study.

After about two years of field research, enough qualitative material had been gathered to permit the research team to develop and test a closed-ended survey instrument that would be administered to all coalition members or a substantial sample of them. Our conceptual frameworks provided excellent guidance about the topics to include in the survey; the years of field research made it possible to specify the likely range of responses to questions. After multiple iterations and expert review, the survey was cognitively tested by being administered to people similar to the expected respondents (in this

case, members of local tobacco control coalitions in states not included in our sample). The findings from initial administration of the survey have been "fed back" to the sites.

The research team is now carrying out an integrative, cross-site analysis of both qualitative and quantitative data; we are also developing briefer case studies designed to highlight key "lessons learned" from the research. The integrative analysis uses key common variables addressed in the qualitative research, typically but not always by creating dichotomous variables; quantitative data from the administration of the survey; and, finally, supplementary secondary data on environmental characteristics (e.g., items from the Area Resource File); changes in state and, to the extent available, local tobacco control policies; and public opinion data drawn from surveys supported by the Foundation through another grant. We believe, however, that our ability to interpret the results of this more quantitative analysis will be substantially enhanced by our experiences in the field.

FINAL THOUGHTS

As noted at the outset, health services and policy research is young as a field of inquiry. At this stage of its development, researchers and funders cannot afford to ignore the potential contributions of qualitative methods in identifying important questions, in building the capacity to conduct and replicate research, and in constructing useful theories. However, the contributions of qualitative research will not be maximized unless the methods are applied with rigor as well as creativity. The purchasers and consumers of research have every right to demand vigilance in ensuring that those who design and conduct this research have the right training and experience and that they acknowledge and, to the extent possible, protect against the investigator-dependent nature of this research.

NOTES

1. Indeed, as described in one of our examples, qualitative research is particularly useful in dealing with unanticipated events. Luft (1986) has also described a mixed-method study in which qualitative research findings not only helped point to significant problems and anomalies in the secondary, quantitative database his research team had planned to use, but proved to be the only truly reliable data source to address their questions.

2. Purposive sampling is often combined with "snowball" sampling, in which additional informants or events that deserve investigation are identified as you go along. Note that purposive sampling can also be designed to look not only for "typical" cases, but for atypical or extreme cases or for cases that apparently do not fit the underlying assumptions or hypotheses of the research. Thus, it can be designed specifically to reduce the potential for bias and support, again, the need to look for "disconfirmation" of emergent hypotheses.

3. Ironically, the conceptual framework was originally designed to guide an assessment of coalitions set up under the National Cancer Institute's quite similar tobacco control initiative, the ASSIST project. The qualitative nature of this research, however, contributed to the decision that such an assessment would not be carried out.

REFERENCES

Babbie, E. 1998. *The Practice of Social Research, 8th ed.* Belmont, CA: Wadsworth Publishing Company.

Bowling, A. 1997. *Investigating Health and Health Services.* Briston, PA: Open University Press.

Campbell, D. T. 1975. "Qualitative Knowing in Action Research." In *The Social Context of Methods,* edited by M. Brenner, P. Marsh, and M. Brenner, pp. 184–209. London: Croom Helm.

Caudle, S. L. 1994. "Using Qualitative Approaches." In *Handbook of Practical Program Evaluation,* edited by J. S. Wholey, M. P. Hatry, and K. E. Newcomer, ch. 4. San Francisco: Jossey-Bass.

Corbin, J. M., and A. L. Strauss. 1988. *Unending Work and Care: Managing Chronic Illness at Home.* San Francisco: Jossey-Bass.

Cronbach, L. J. 1963. "Course Improvement Through Evaluation." *Teachers College Record* 64 (5): 672–83.

Diamond, T. 1992. *Making Gray Gold: Narratives from Inside the Nursing Home.* Chicago: University of Chicago Press.

Glaser, B. 1965. *Awareness of Dying.* Chicago: Aldine.

Glaser, B., and A. Strauss. 1967. *The Discovery of Grounded Theory.* Chicago: Aldine.

Goffman, E. 1961. *Asylums: Essays on the Social Situation of Mental Patients and Other Inmates.* Garden City, NJ: Anchor Books.

Kaplan, A. 1964. *The Conduct of Inquiry.* San Francisco: Chandler Publishing Company.

Lincoln, Y. S., and E. G. Guba, E. G. 1985. *Naturalistic Inquiry.* Beverly Hills, CA: Sage Publications.

Luft, H. S. 1986. "Health Services Research as a Scientific Process: The Metamorphosis of an Empirical Research Project from Grant Proposal to Final Report." *Health Services Research* 21 (4): 563–84.

Miles, M. B., and A. M. Huberman. 1994. *Qualitative Data Analysis: An Expanded Sourcebook.* Beverly Hills, CA: Sage Publications.

Morgan, D. L. 1988. *Focus Groups as Qualitative Research.* Newbury Park, CA: Sage Publications.

Patton, M. Q. 1997. *Utilization-Focused Evaluation, 3d ed.* Beverly Hills, CA: Sage Publications.

———. 1990. *Qualitative Evaluation and Research Methods, 2d ed.* Newbury Park, CA: Sage Publications.

———. 1975. *Alternative Evaluation Research Paradigm.* Grand Forks, ND: University of North Dakota.

Ragin, C. C. 1994. *Constructing Social Research: The Unity and Diversity of Method.* Thousand Oaks, CA: Pine Forge Press.

Schaps, E., S. Churgin, and Research Team. 1981. "A Review of 127 Drug Abuse Prevention Programs." *Journal of Drug Issues* (Special Issue, winter).

Scheirer, M. A. 1994. "Designing and Using Process Evaluation." In *Handbook of Practical Program Evaluation,* edited by J. S. Wholey, H. P. Hatry, and K. E. Newcomer, ch. 3. San Francisco: Jossey-Bass.

Scriven, M. 1967. "The Methodology of Evaluation." In *Perspectives of Curriculum Evaluation,* edited by R. W. Tyler, R. M. Gagne, and M. Scriven. Chicago: Rand McNally.

Strauss, A. L. 1978. *Negotiations: Varieties, Contexts, Processes and Social Order.* San Francisco: Jossey-Bass.

Wholey, J. S., H. P. Hatry, and K. E. Newcomer, eds. 1994. *Handbook of Practical Program Evaluation.* San Francisco: Jossey-Bass.

Yin, R. K. 1984. *Case Study Research: Design and Methods.* Beverly Hills, CA: Sage Publications.